Some Performance Limits in Imaging and Image Processing

Peyman Milanfar*
EE Department
University of California, Santa Cruz
milanfar@ee.ucsc.edu
http://www.soe.ucsc.edu/~milanfar

*Joint work with Dirk Robinson, Morteza Shahram, and Sina Farsiu, Michael Elad, Ali Shakouri

Motivation

- “The main focus of the workshop will be the analysis of image or image-like data with a view to the rigorous analysis of data from scientific experiments.”
  - Estimation of Motion
  - Resolution Enhancement
  - Edge Detection

Topic I: Motion Estimation

Motion Estimation: Where are we?

"After some 20 years work on motion estimation, I think we know what we're doing"

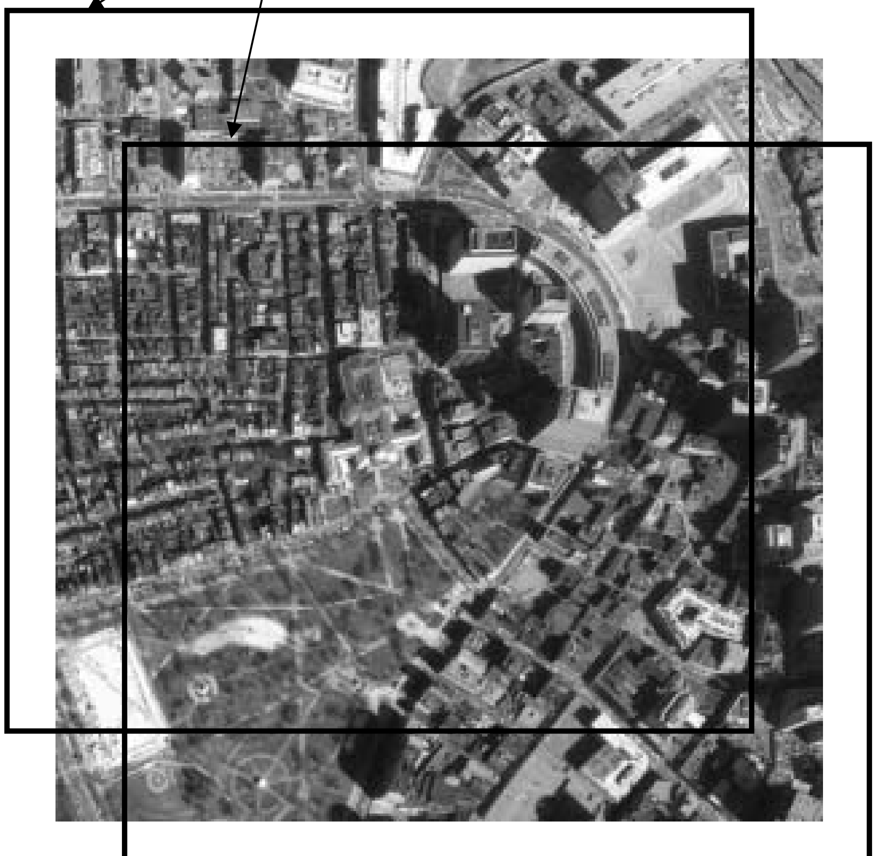
-David Fleet (PARC)

Motion Estimation: A Model

- Signal Model

$$f_1(x,y) = f(x,y) + n_1(x,y)$$

$$f_2(x,y) = f(x-v_x, y-v_y) + n_2(x,y)$$



Milanfar et al. EE Dept, UCSC

Statistical Solution: Maximum Likelihood

- Correlation methods

$$\max_{v_x, v_y} \left(\sum_{x,y} f_1(x - v_x, y - v_y) f_2(x, y) \right)$$

- Nonlinear Least Square:

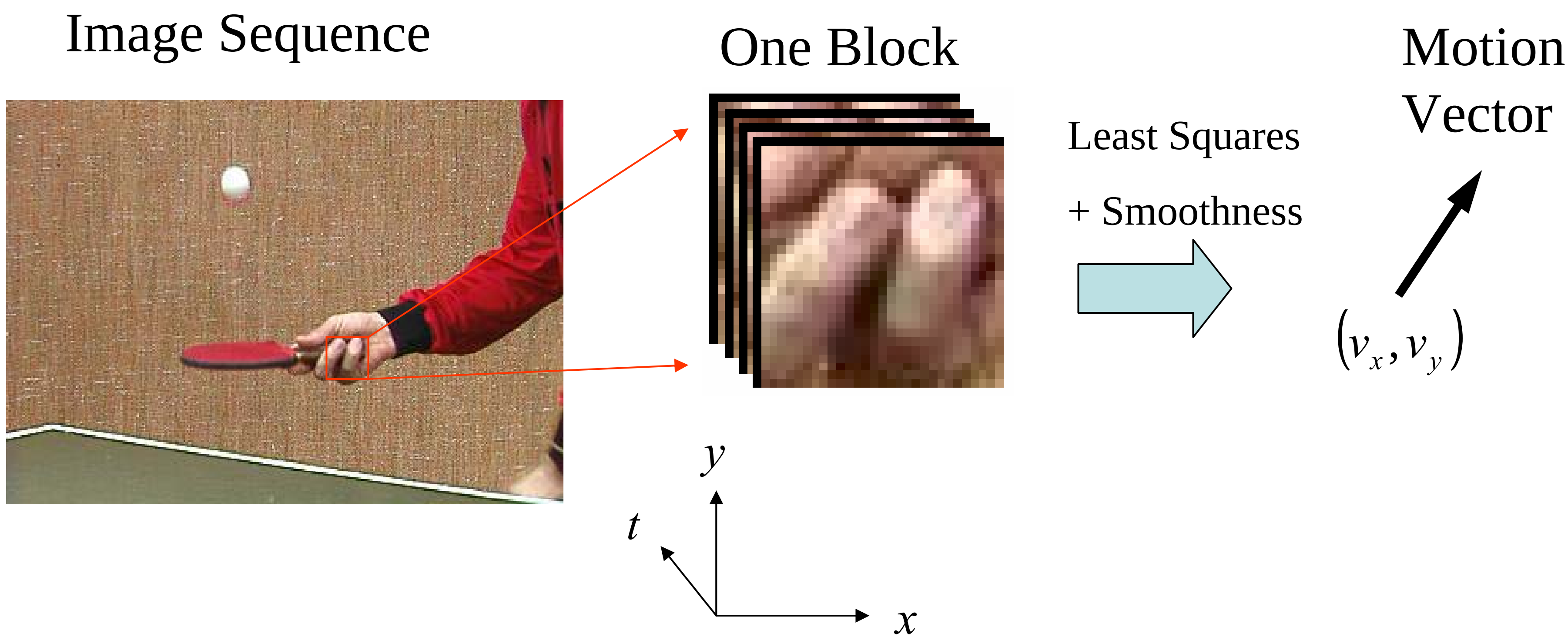
$$\min_{v_x, v_y} \sum_{x,y} \left(f_1(x - v_x, y - v_y) - f_2(x, y) \right)^2$$

- Improving to subpixel acuracy:
  - Fits a quadratic about the peak of the correlation surface.
  - Gauss-Newton methods, iterated improvement,
    - iterating over scale: pyramid-based methods

The Optical Flow Method

- Optical Flow Equation (Linear Least Squares)

$$\frac{d}{dt} f(x,y,t) = f_x \cdot v_x + f_y \cdot v_y + f_t = 0 \quad (\approx \nabla f^T v + f_1 - f_2)$$



Milanfar et al. EE Dept, UCSC

Performance Limits

- Fisher Information Matrix (Assume GWN)

$$I(v_x, v_y) = \frac{1}{\sigma^2} \begin{bmatrix} \sum_{x,y} f_x^2(x - v_x, y - v_y) & \sum_{x,y} f_x f_y(x - v_x, y - v_y) \\ \sum_{x,y} f_x f_y(x - v_x, y - v_y) & \sum_{x,y} f_y^2(x - v_x, y - v_y) \end{bmatrix}$$

- Bound on the mean-squared error

$$Q = E[(v - \hat{v})(v - \hat{v})^T] \geq I^{-1}(v_x, v_y) = J$$

$$Q_{11} = E[(v_x - \hat{v}_x)^2] \geq J_{11}$$

$$Q_{22} = E[(v_y - \hat{v}_y)^2] \geq J_{22}$$

How close to the limit?

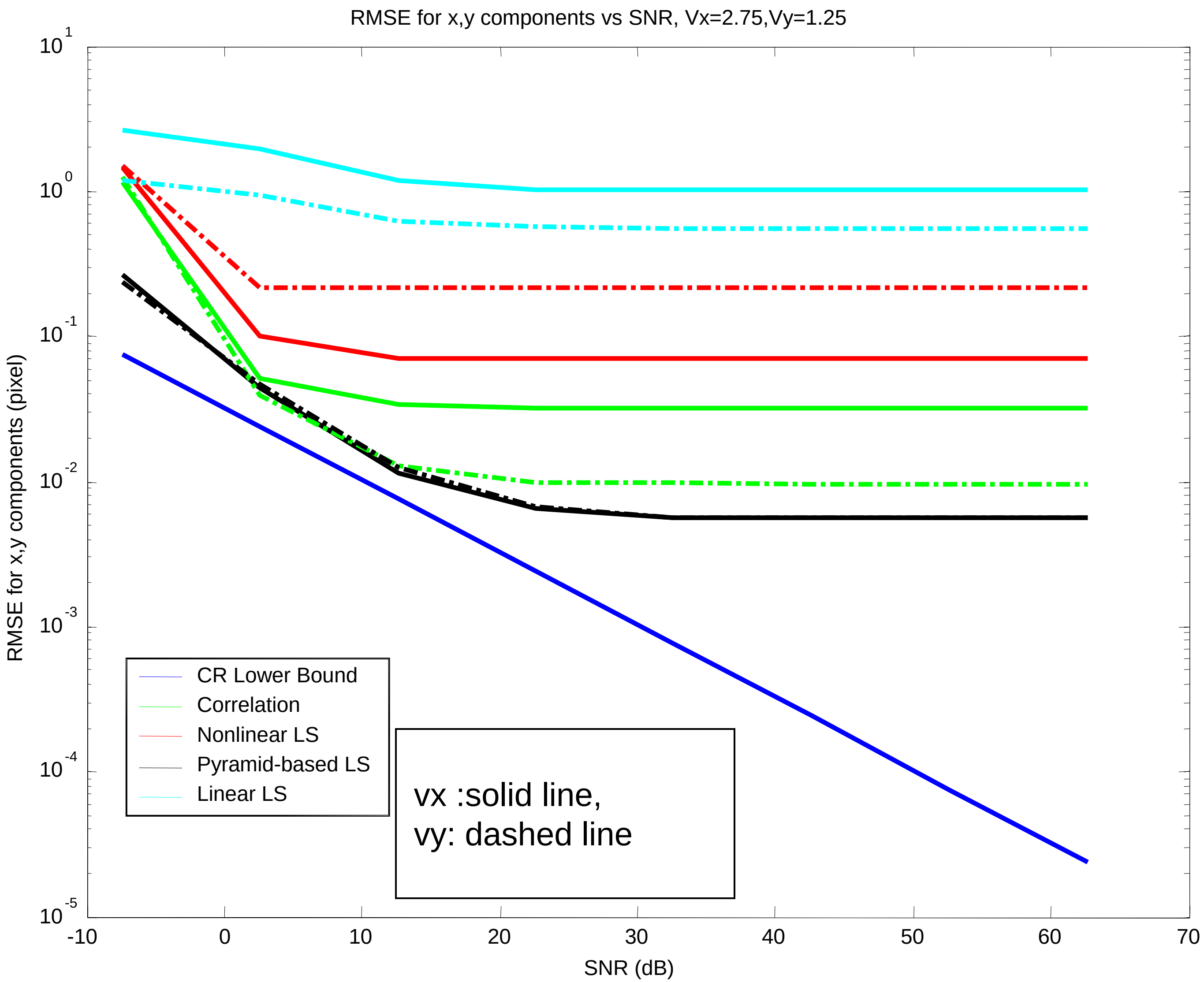
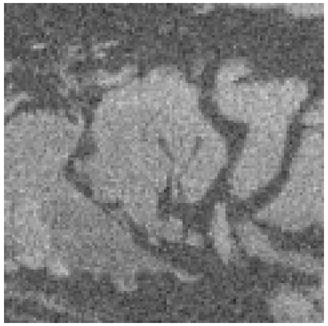


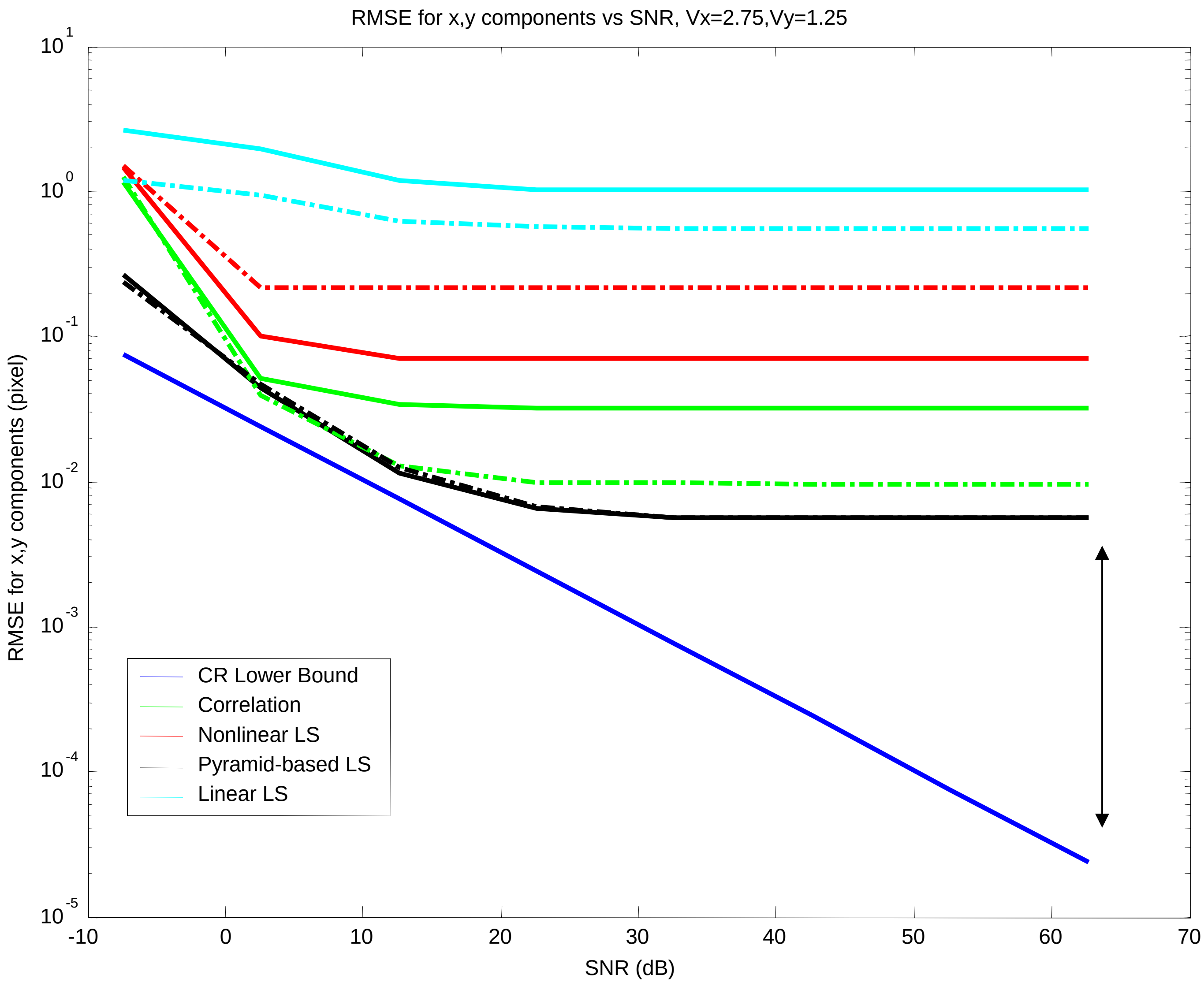
Image used:

At 3 dB:



et al. EE Dept, UCSC

What happens at high SNR?



Bias!

Performance vs Image content

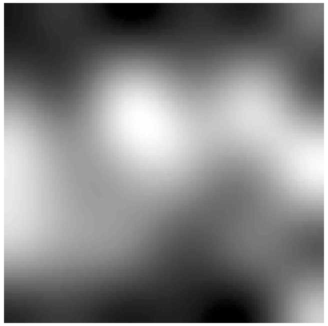
face

forest

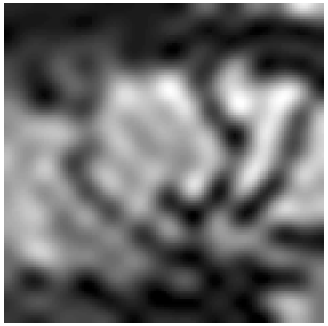
lab

tree

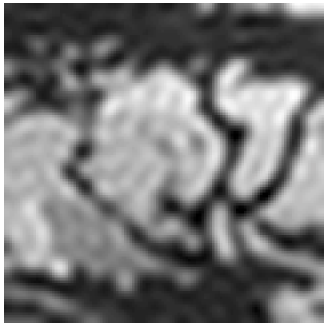
Image as % of Full Bandwidth



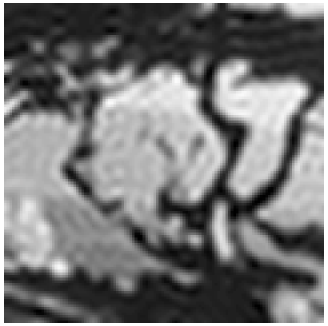
4%



12%



20%

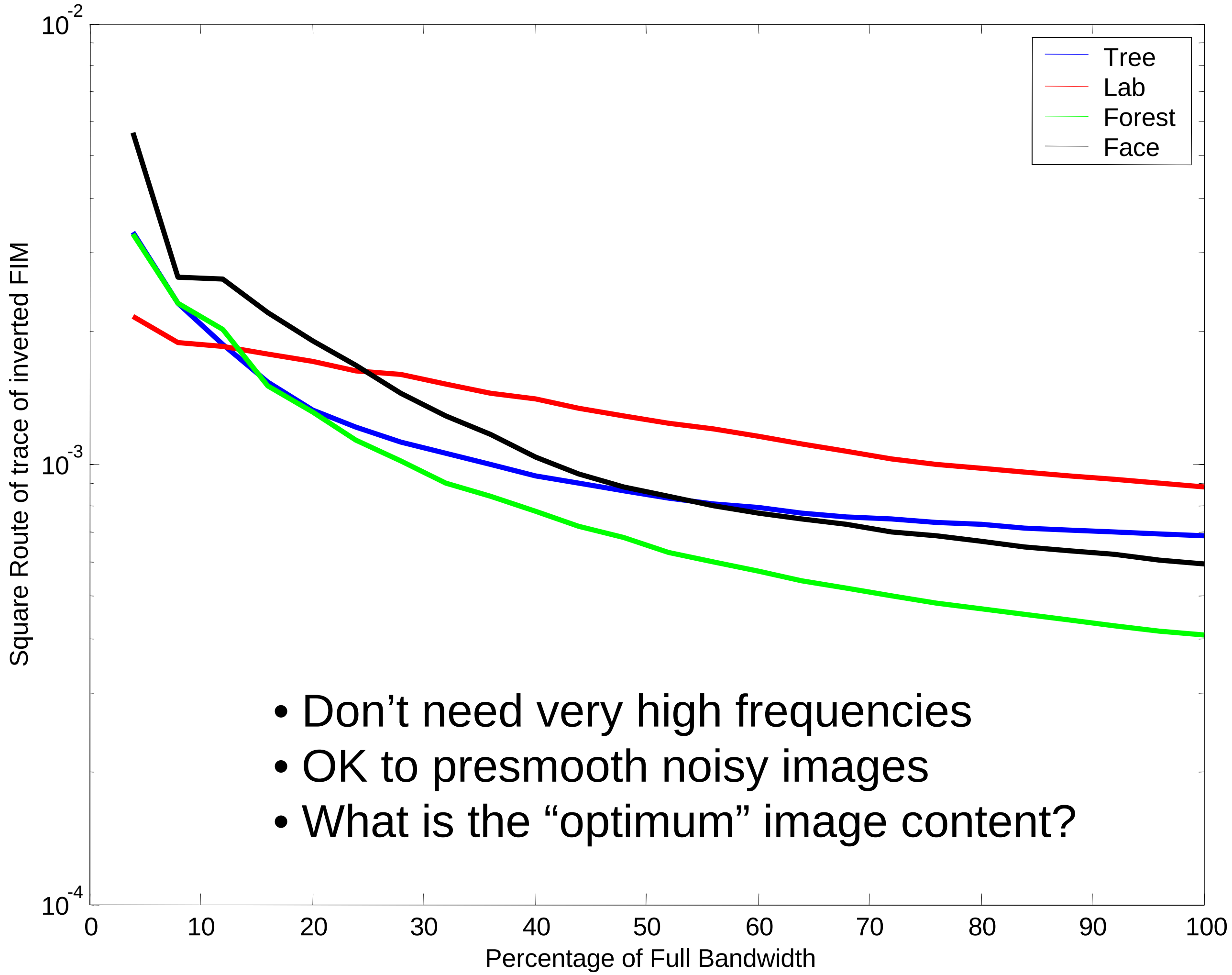


28%

Milanfar et al. EE Dept, UCSC

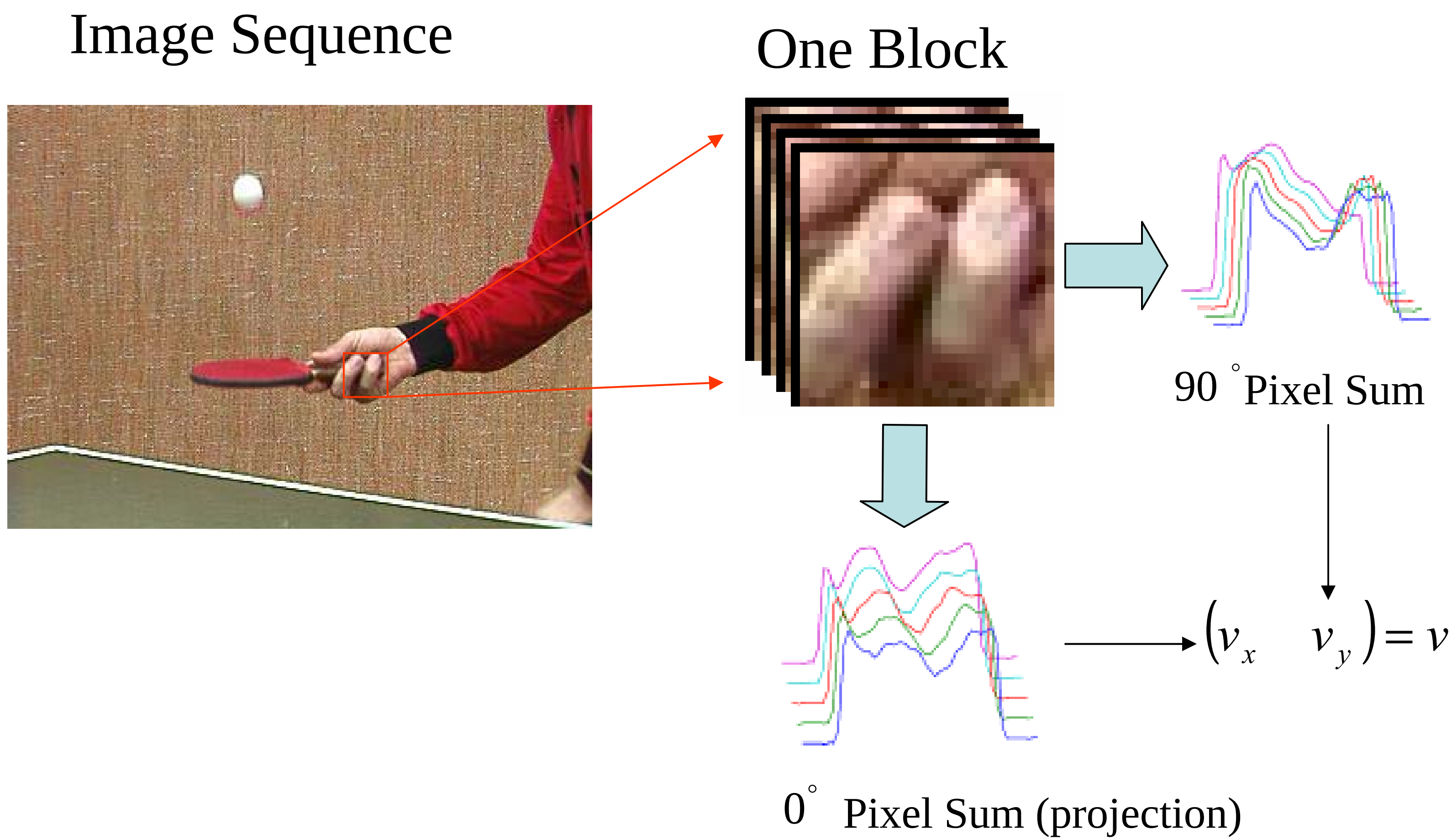
Performance vs Image content

Sqrt of Trace of J



- Don’t need very high frequencies
- OK to presmooth noisy images
- What is the “optimum” image content?

Can the limitations be overcome?



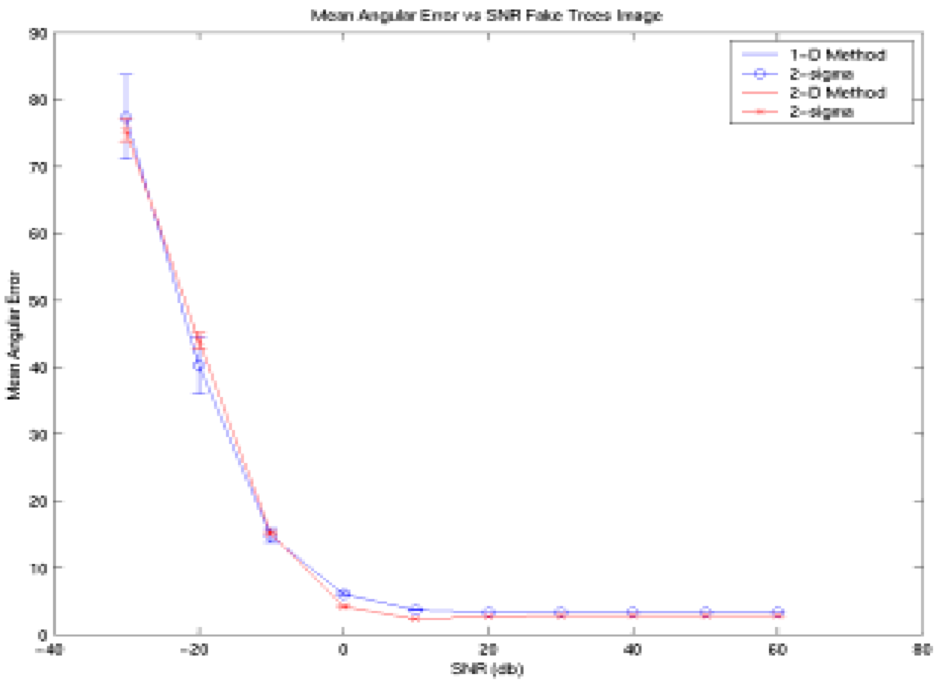
Milanfar et al. EE Dept, UCSC

Qualitative Comparison

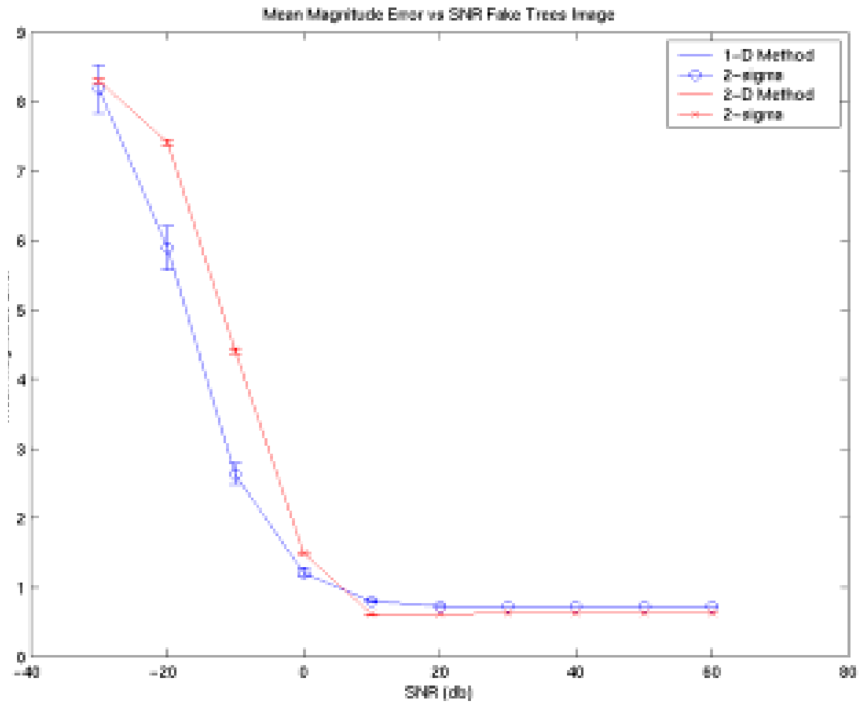
- Performance:
  - In most cases, performance of 2-D and 1-D methods are within 5% of each other (mean magnitude or angular error)
- Complexity:
  - Projection-based an order of magnitude faster
- Surprising fact:
  - Projection can *improve* performance.

Quantitative Comparison

Mean angular error vs. SNR



Mean magnitude err. vs. SNR

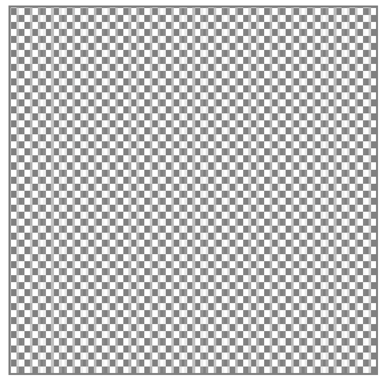


1-D Method better at low SNRs

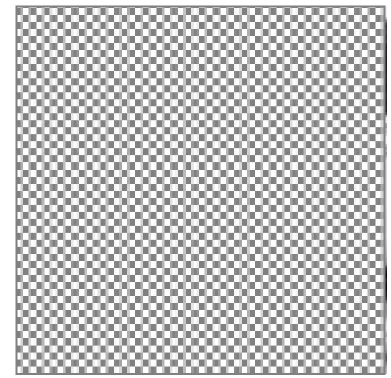
Why Improvement?

- Interference Rejection

+



=



Moving

Still or known motion

Sum

“Same” row/column projections

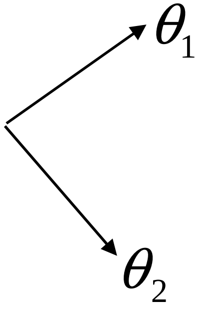
Why Improvement?

- 2-D Spectrum vs. 1-D Spectrum
  - Projection - Slice Theorem:
    - f → 2-D Fourier Transform → Slice @ angle θ
    - f → Projection sum at θ → 1-D Fourier Trans.
- Smoothness
  - Radon transform of f is ½ degree smoother than f.
- Reduced Bias

Optimal Projection Angles

- Goal: Find the set of projection angles that minimizes mean-square error.

- Partial solution: Find a pair of orthogonal directions where the product of "power" in the derivative of projections is largest. (What about the bias?)



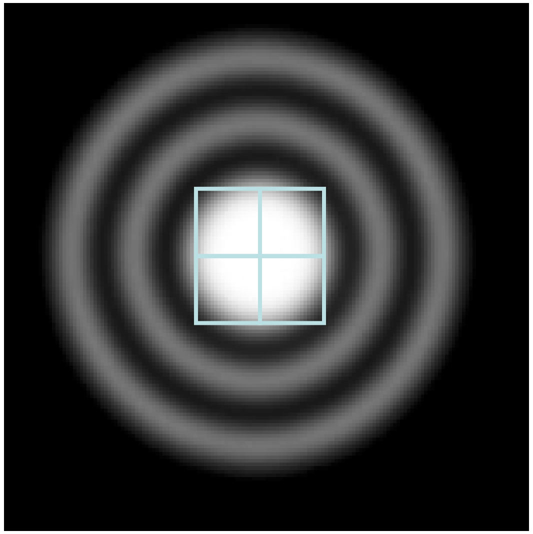
Topic I: Summary

- Limits to how well motion can be estimated.
  - Existing methods should be measure against these limits.
- Bias-variance tradeoff in motion estimation. What is best?
  - Bias is hard to characterize
- What are “best” image patterns for motion estimation?

Topic II: Resolution Enhancement

Why Resolution Enhancement?

- To obtain an alias-free, “diffraction limited” image we need 4 pixels covering the Airy disk:



- That is: radius of the Airy disk must match the pixel dimensions.

Resolution Enhancement Idea

- Given multiple low-resolution moving images of a scene (a video), generate a high resolution image (or video).

$$\underbrace{\text{frame}_1, \text{frame}_2, \cdots, \text{frame}_{N\text{-}1}, \text{frame}_N}_{\text{High Resolution Frame}}$$

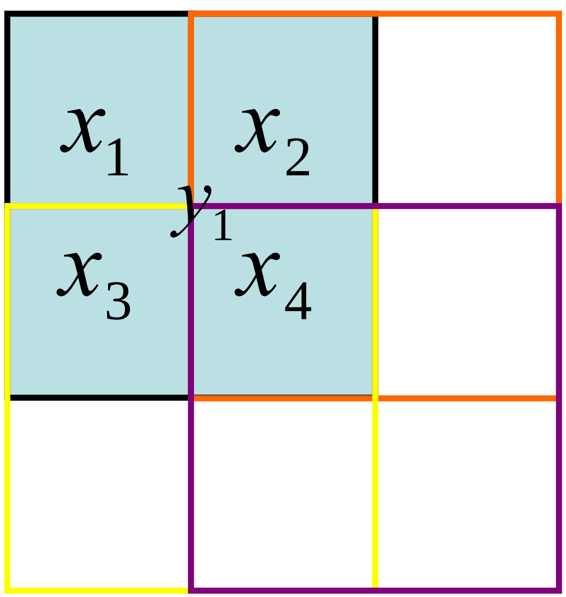
$$\underbrace{\text{frame}_1, \text{frame}_2, \cdots,}_{\text{High Resolution Frame}_1} \; \underbrace{\cdots, \text{frame}_{N\text{-}1}, \text{frame}_N}_{\text{High Resolution Frame}_2}$$

"Trading off time resolution or view diversity to gain spatial resolution"

Milanfar et al. EE Dept, UCSC

Resolution Enhancement Model

- A simple model relating the low-resolution blurry image to the high resolution crisper image.



“PSF”

$$y_1 = a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + e_1$$

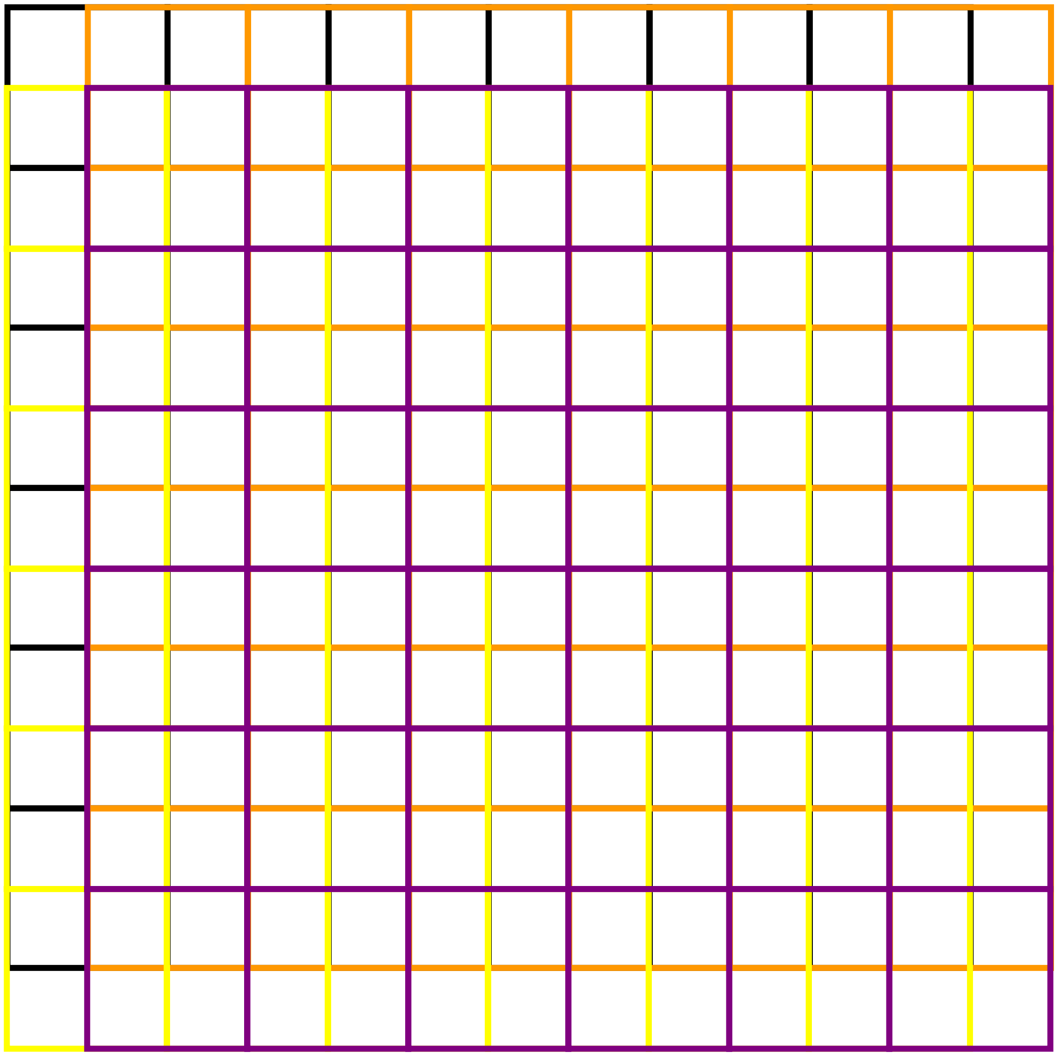
$$y_2 = 0 \cdot x_1 + a_2 x_2 + 0 \cdot x_3 + a_4 x_4 + e_2$$

$$y_3 = 0 \cdot x_1 + 0 \cdot x_2 + a_3 x_3 + a_4 x_4 + e_3$$

$$y_4 = 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + a_4 x_4 + e_4$$

Low vs High Res Pixels

x2 enhancement



Milanfar et al. EE Dept, UCSC

The Mathematical Model

$$\underbrace{\underline{y}_k}_{\text{k-th frame}} = A_k \underline{x} + \underline{e}_k \qquad for \qquad 1 \le k \le p$$

$$A_k = D C_k W_k$$

Downsampling — D, Blurring — C_k, Warping — W_k

$$A_k = [T_{k,1} \quad T_{k,2} \quad \cdots \quad T_{k,l^2}]$$

Upper-banded, “nearly” Toeplitz — T_{k,l^2}

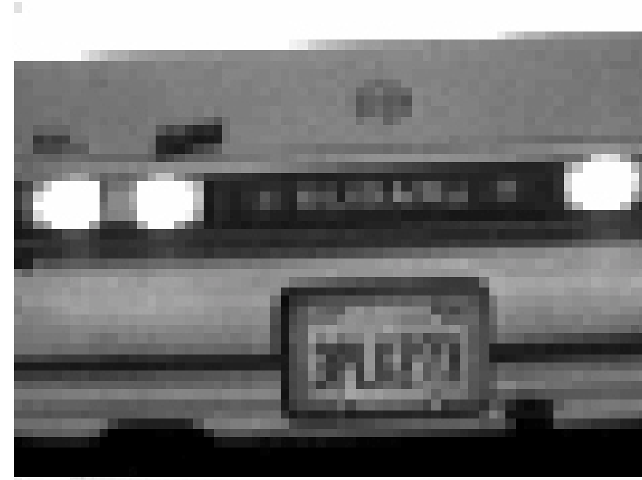
BTTB system → $$y = Ax + e$$

- The system is typically underdetermined and ill-conditioned.
- 10’s or 100’s of thousands of unknown variables and data.
- **Warping (motion), must be estimated!**

Some real examples:

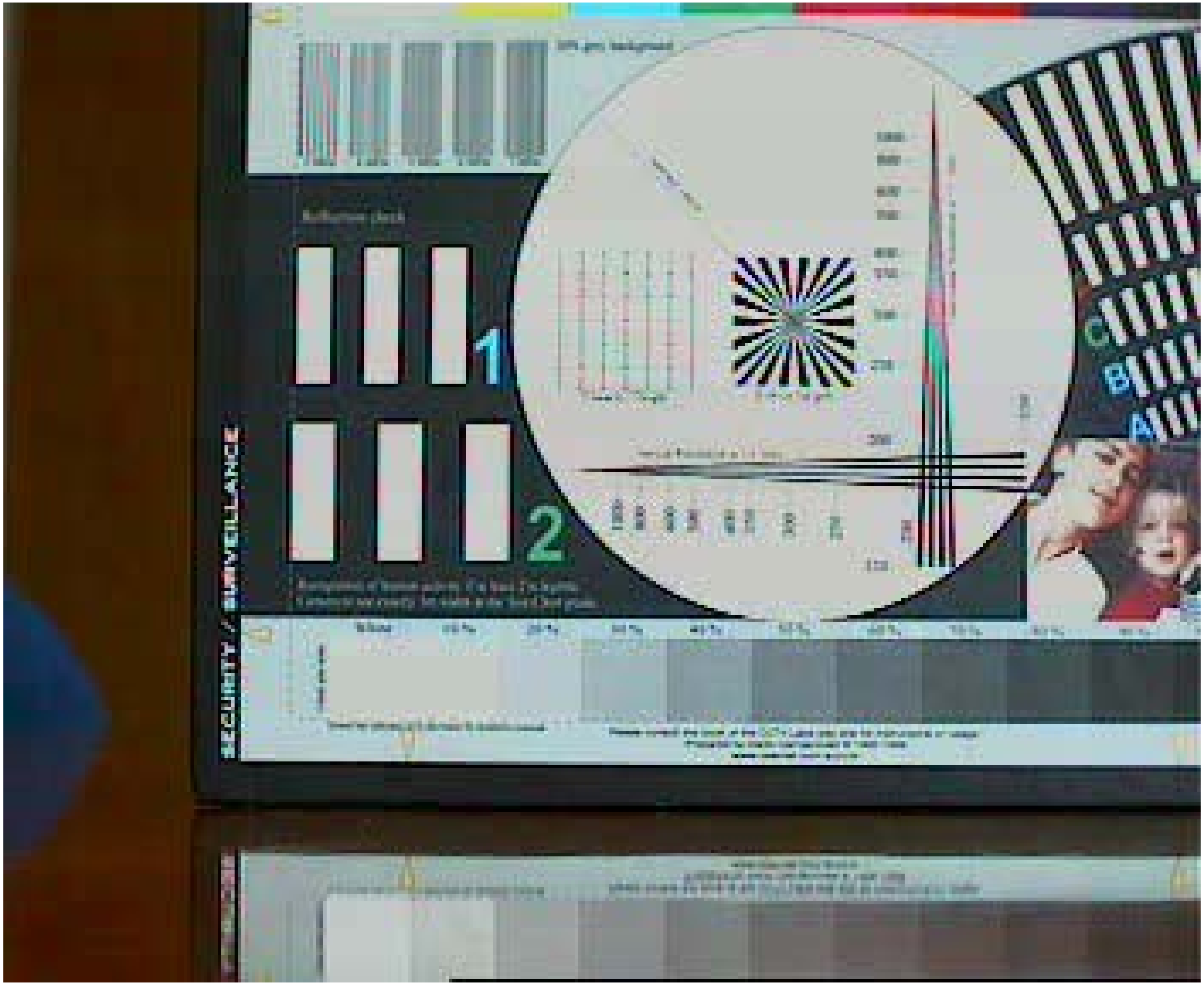
Infrared Camera (Night Vision)

License Plate Reading



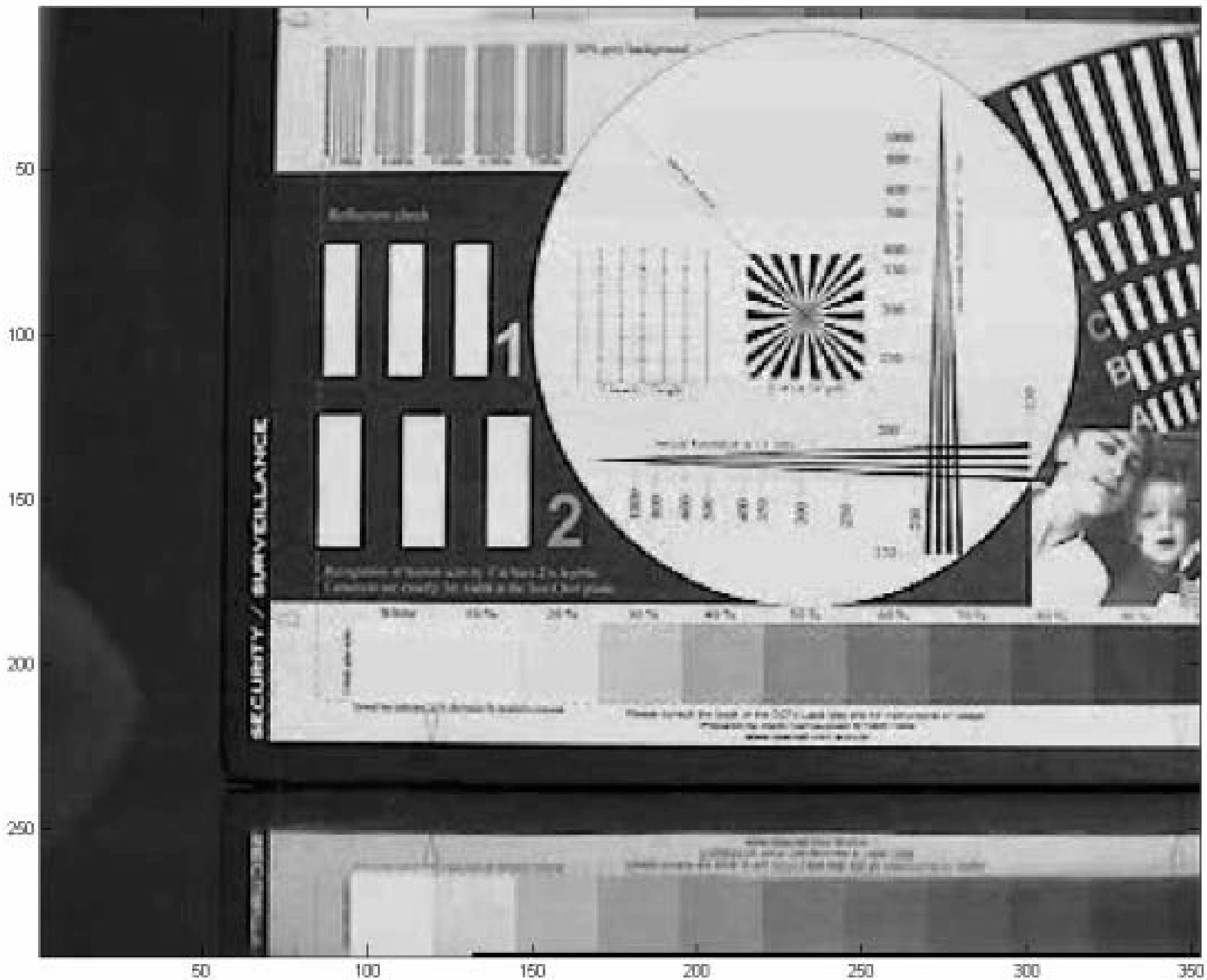
Digital Video Camera from 2nd story window

Milanfar et al. EE Dept, UCSC



Data Courtesy Vigilant Technology

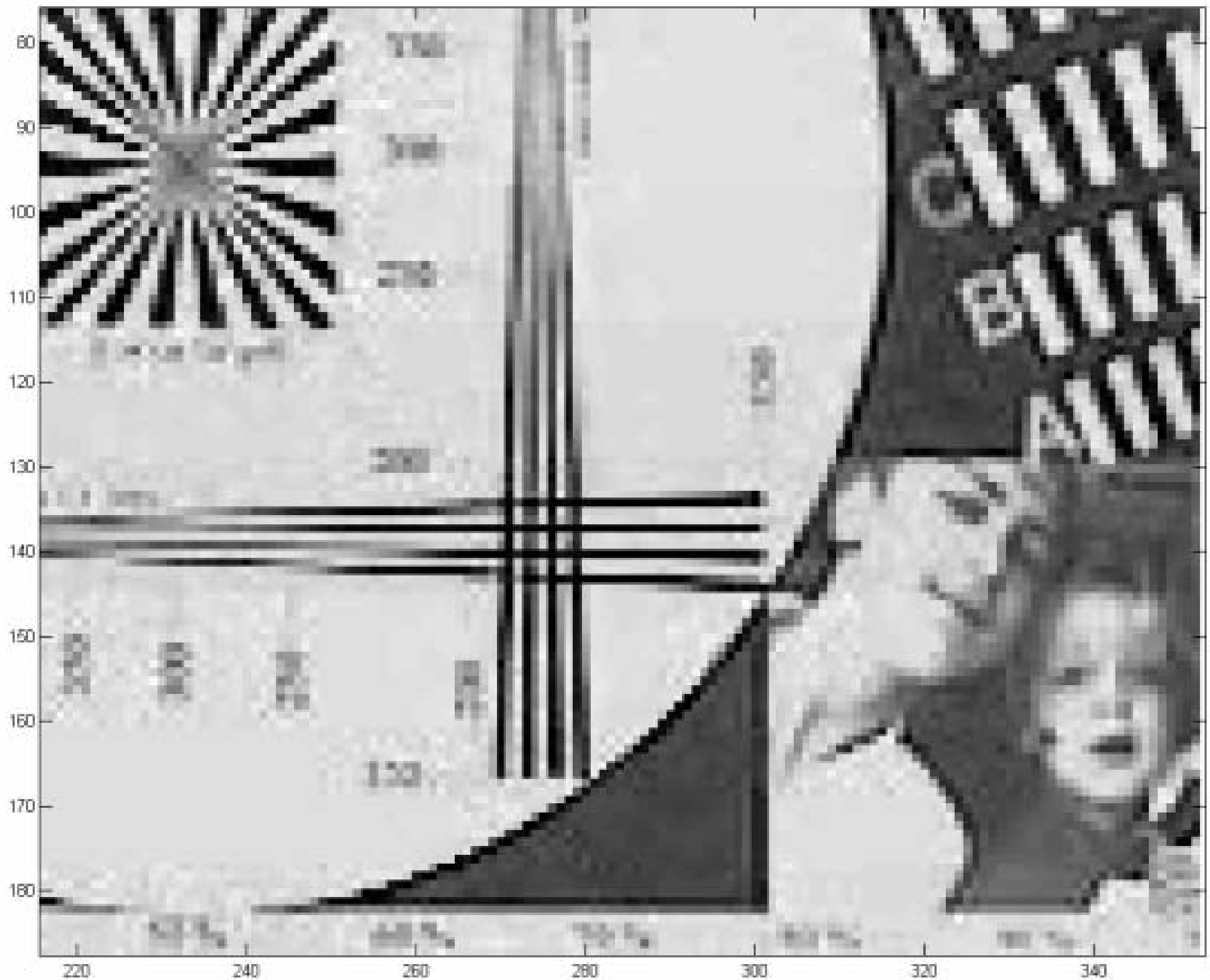
Before


50
100
150
200
250
50
100
150
200
250
300
350
1
2
C
B
A
SECURITY / SURVEILLANCE

After

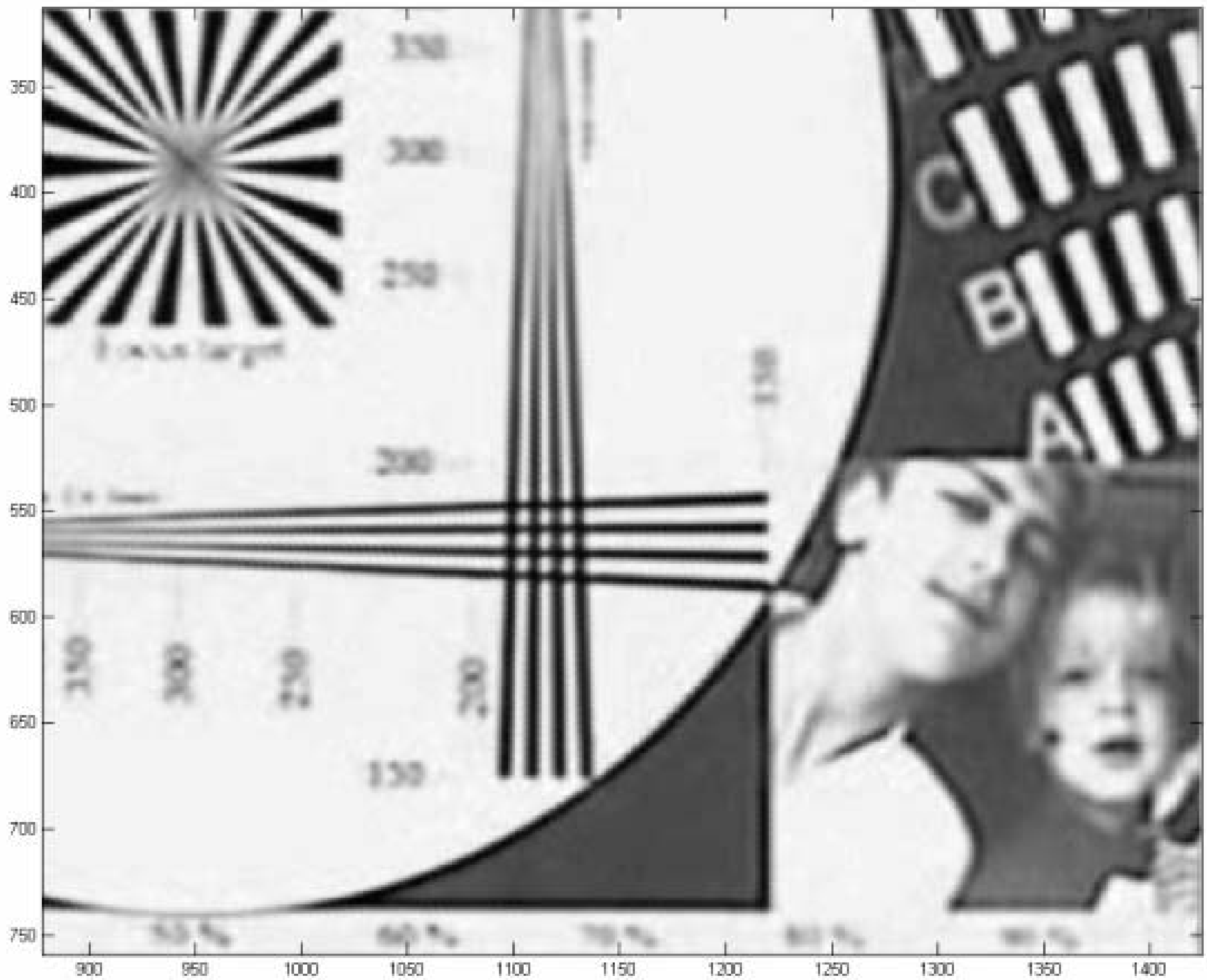

100
200
300
400
500
600
700
800
900
1000
1100
200
400
600
800
1000
1200
1400
1
2
SECURITY / SURVEILLANCE
C
B
A

Detail Before



Data Courtesy Vigilant Technology

Detail After

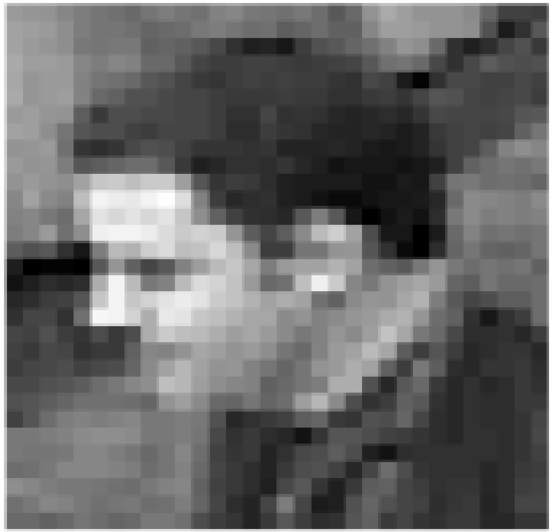


Data Courtesy Vigilant Technology

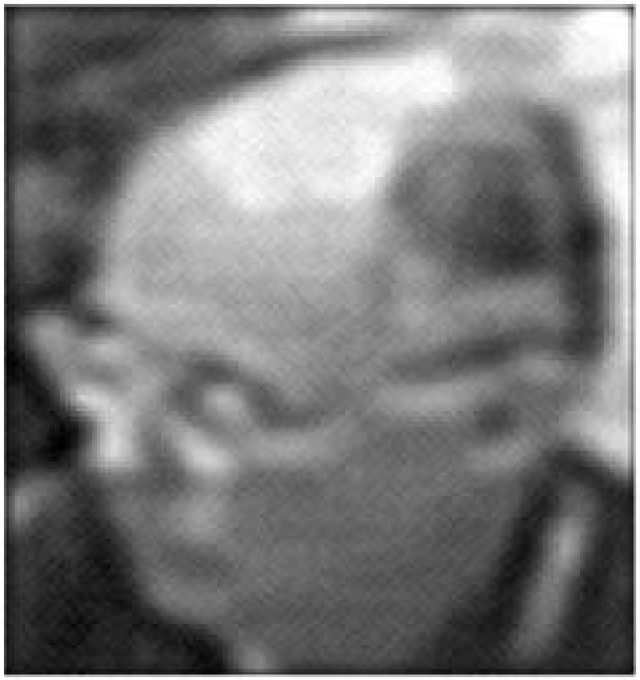
MPEG Surveillance Video

Data Courtesy Vigilant Technology

Milanfar et al. EE Dept, UCSC



Data Courtesy Vigilant Technology



What are the limits to enhancement?

- Motion Estimation Accuracy
- Model Accuracy
- **Sensor Noise**
- **Rayleigh Limit?**

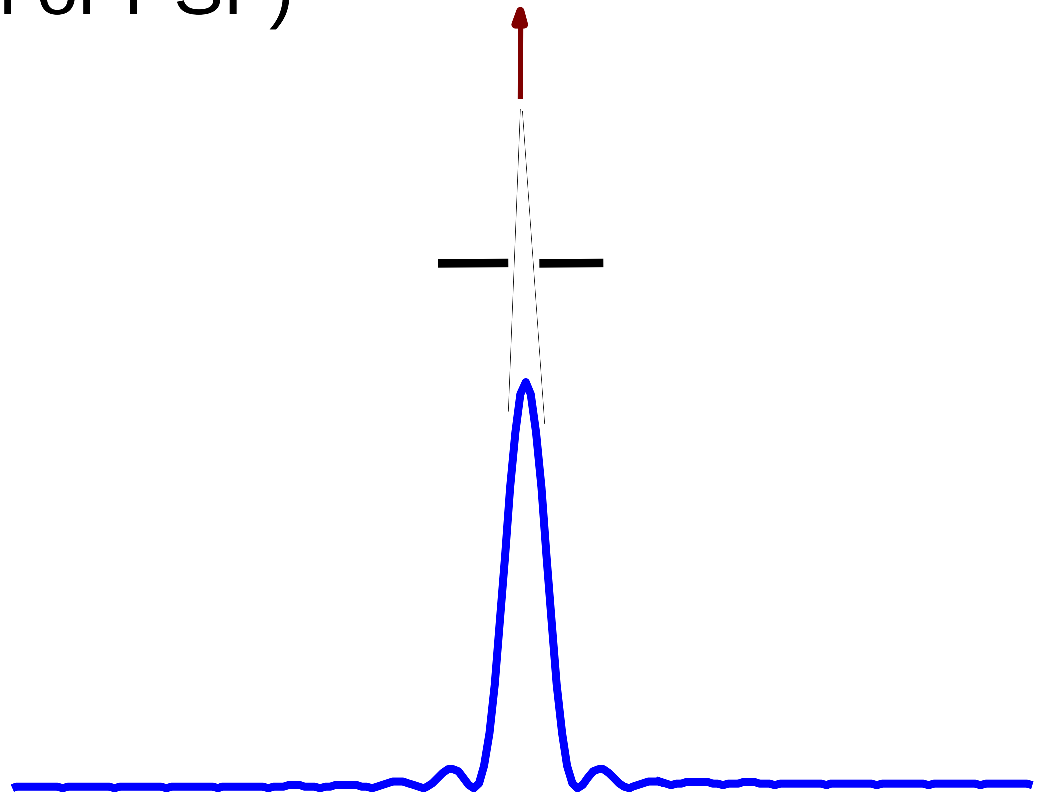
A Case Study

- Incoherent imaging of two closely-spaced point sources

- Statistical definition of resolution: the ability of the imaging system to distinguish two hypotheses in the presence of additive noise.

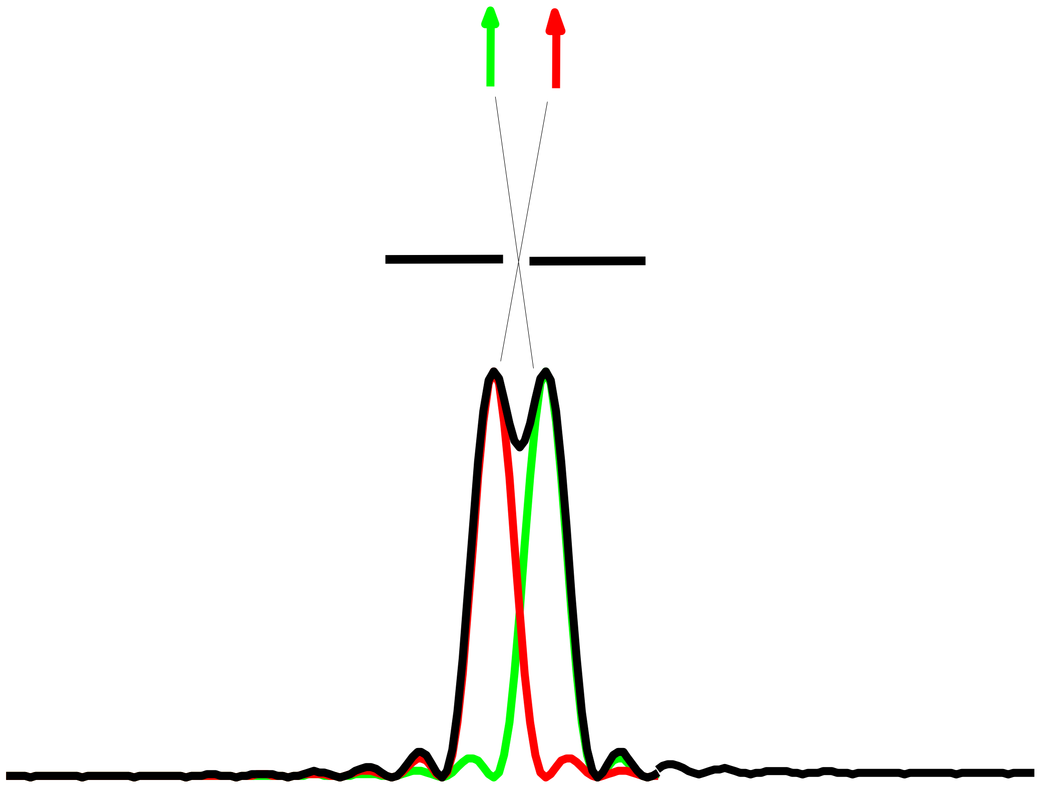
Slit Aperture Imaging

- The image of an ideal point source is captured as a spatially extended pattern (point spread function or PSF)



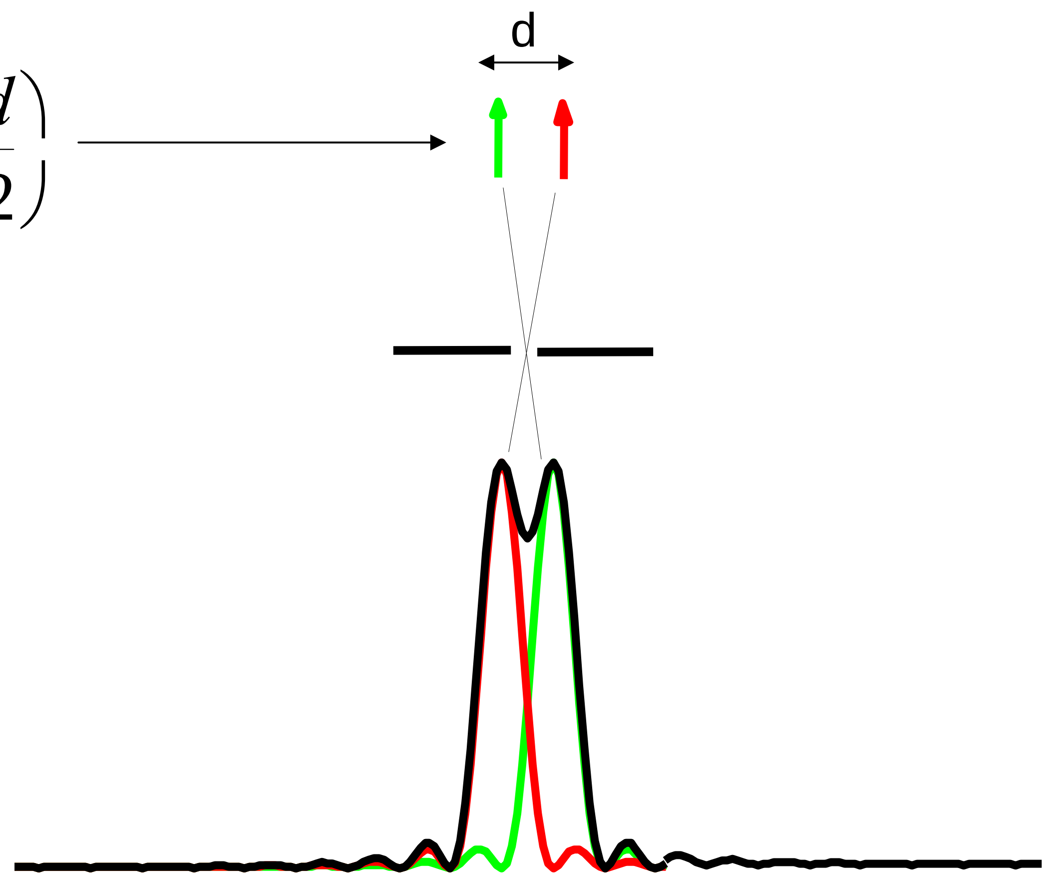
Incoherent Imaging

- The image of two sources is the incoherent sum of PSFs, representing the effect of the diffraction



Incoherent Imaging

$$\sqrt{\alpha}\,\delta\left(x-\frac{d}{2}\right)+\sqrt{\beta}\,\delta\left(x+\frac{d}{2}\right)$$



$$\alpha+\beta=2$$

Power Constraint

GWN(σ^2)

$$f(x,d)=\alpha\,\mathrm{sinc}^2\left(x-\frac{d}{2}\right)+\beta\,\mathrm{sinc}^2\left(x+\frac{d}{2}\right)+w(x)$$

The Rayleigh "Limit"

- According to Rayleigh criterion these two point sources are not resolvable when $d<1$.
  - Rule of thumb, not physical law
- Depending on the signal-to-noise ratio (SNR), resolution beyond the Rayleigh limit is indeed possible. ("super-resolution").
  - But this has its limits too.

A Statistical Definition of Resolution

- The question of presence of one peak (d=0) or two peaks (0<d<1) can be formulated in statistical terms by defining two hypotheses:

$$\begin{cases} H_0 : d = 0 \quad \textit{one peak} \\ H_1 : d > 0 \quad \textit{two peaks} \end{cases}$$

$$\begin{cases} H_0 : f(x) = 2\operatorname{sinc}^2(x) + w(x) \\ H_1 : f(x) = \alpha \operatorname{sinc}^2\left(x - \frac{d}{2}\right) + \beta \operatorname{sinc}^2\left(x + \frac{d}{2}\right) + w(x) \end{cases}$$

Resolution = Discrimination with Unknown Parameters

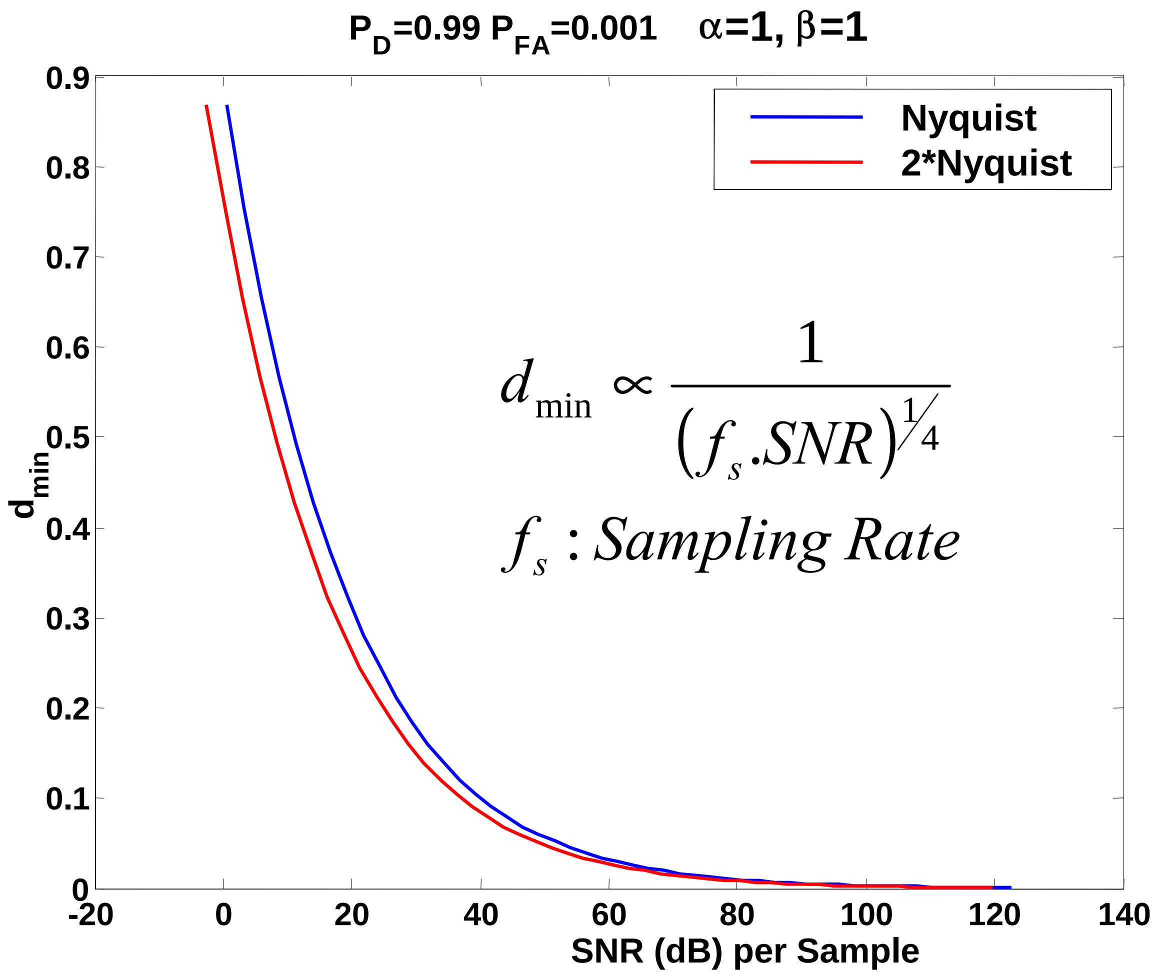
- This is a (nonlinear) problem of signal discrimination with unknown parameter.

- The problem of interest revolves around the values of d in the range $0 \leq d < 1$. We can develop locally optimal statistical tests (discriminators).

Definition of Resolution Limit

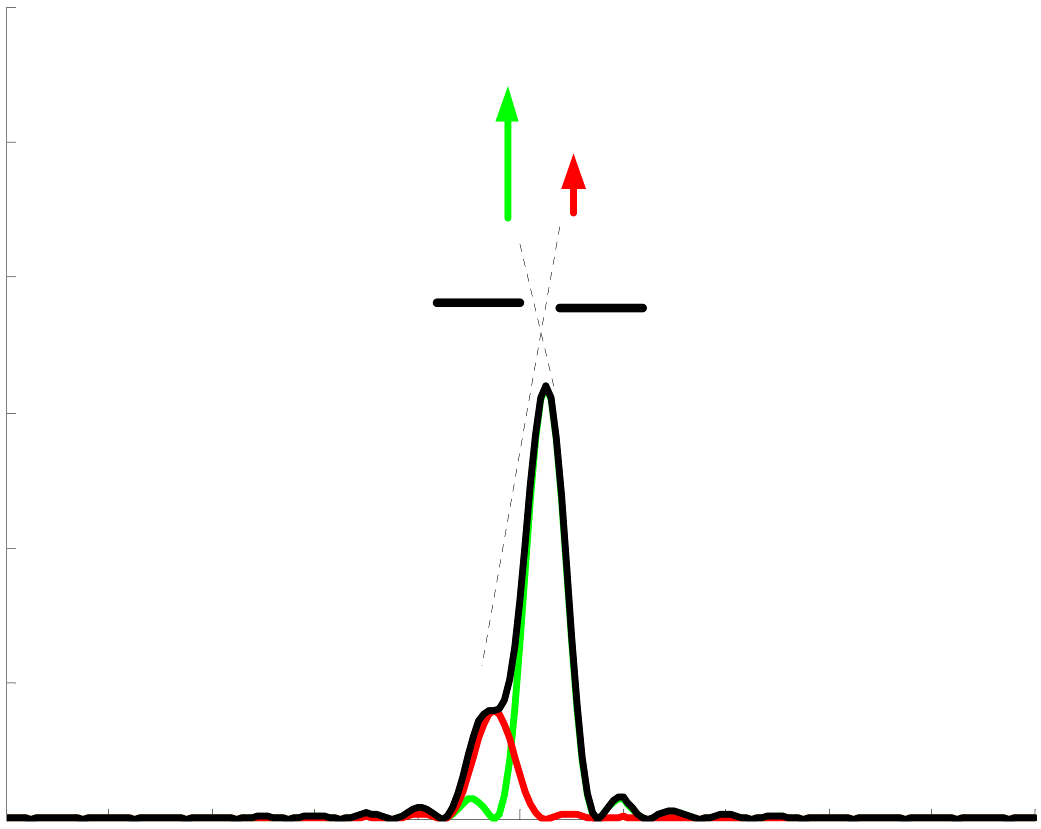
Question: Minimum "d" detectable at very high probability of detection (P_d=0.99) and very low false alarm rate (P_f=0.001) at a given SNR.

Minimum Detectable "d" vs. SNR (equal power)

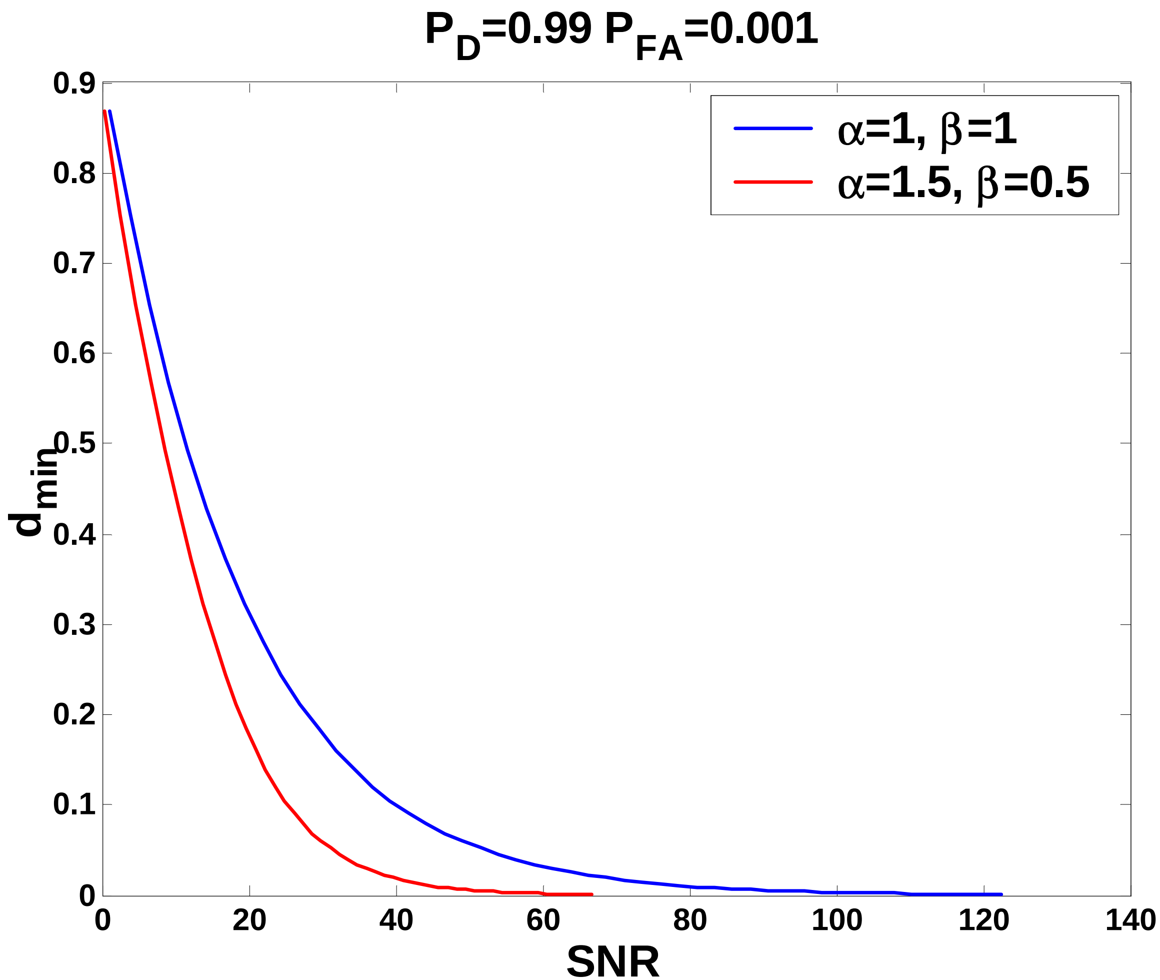


Milanfar et al. EE Dept, UCSC

General Case $\alpha \neq \beta$



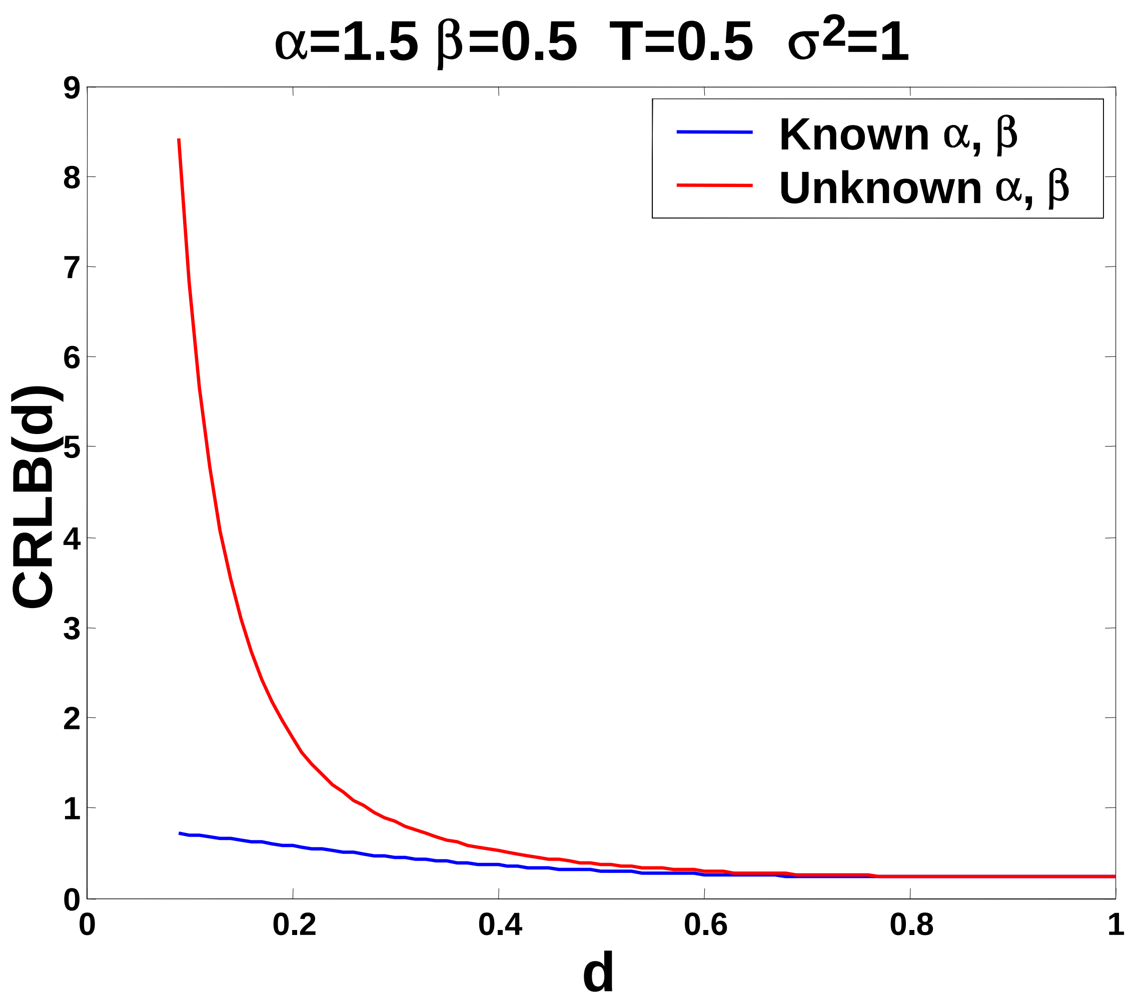
Minimum Detectable "d" vs. SNR (unequal vs. equal power)



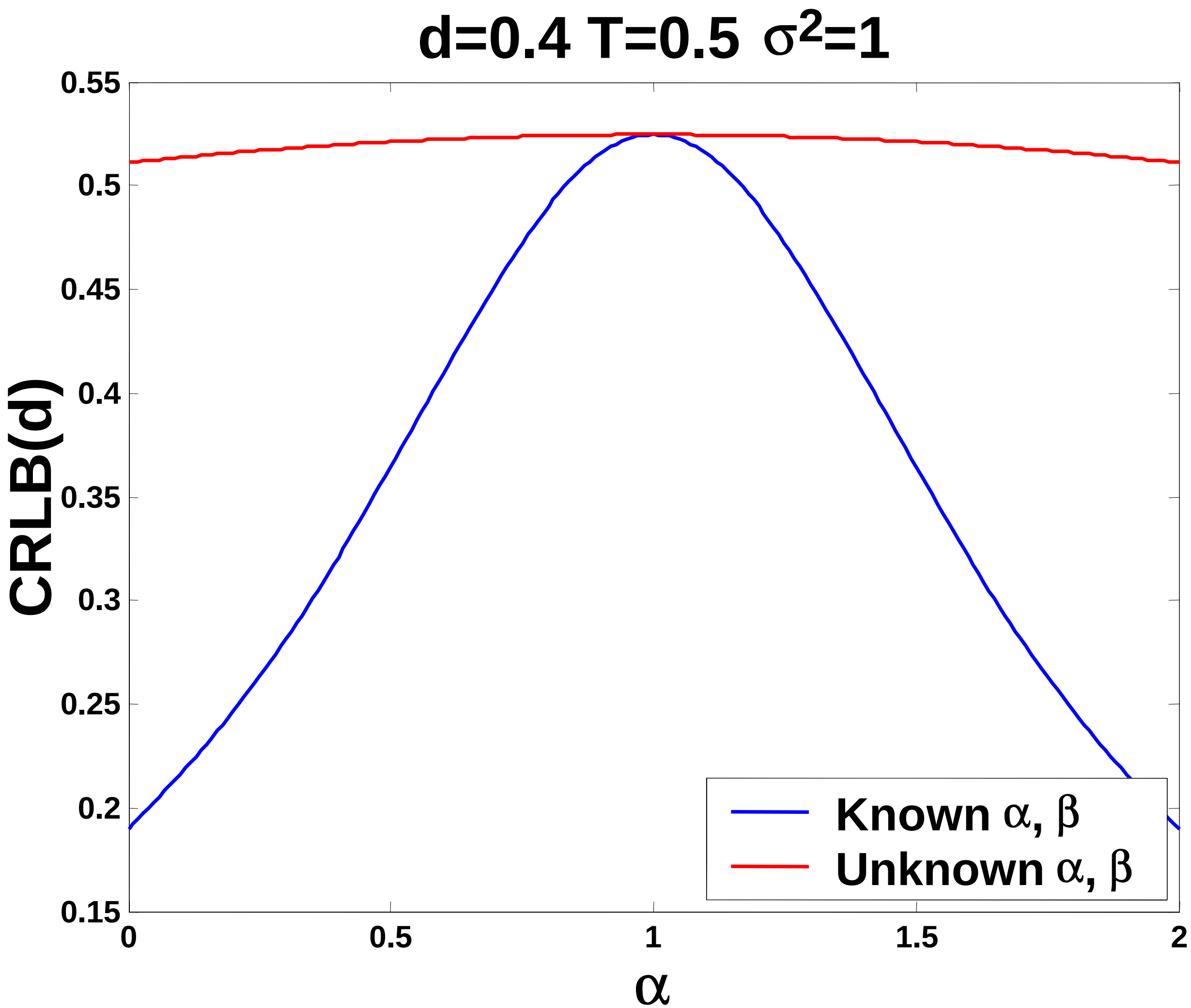
An explanation and some insight

- For a given d_{min}, lower SNR is needed for the case $\beta \neq \alpha$ as compared to the case $\beta = \alpha$.
- Does the information content of the estimate of d behave this way?
  - Yes.

CRLB Curves for d



CRLB Curves for d



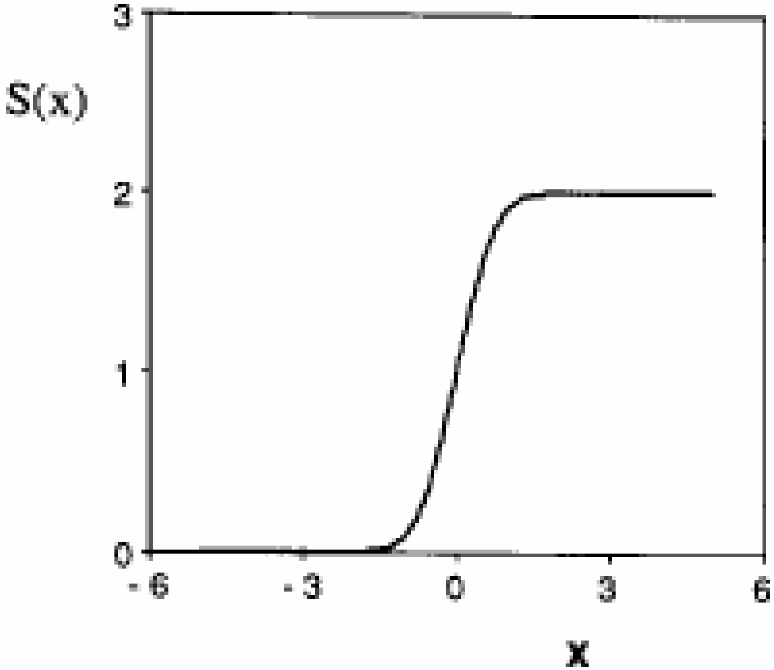
Topic II:Summary

- Two factors play key roles in determining resolution beyond the Rayleigh limit
  - SNR per sample
  - Sampling Rate

- We can address the question of resolution in the context of information theory.

- Extensions to the full resolution enhancement problem, including motion uncertainty, remain to be studied.

Achievable accuracy in edge localization

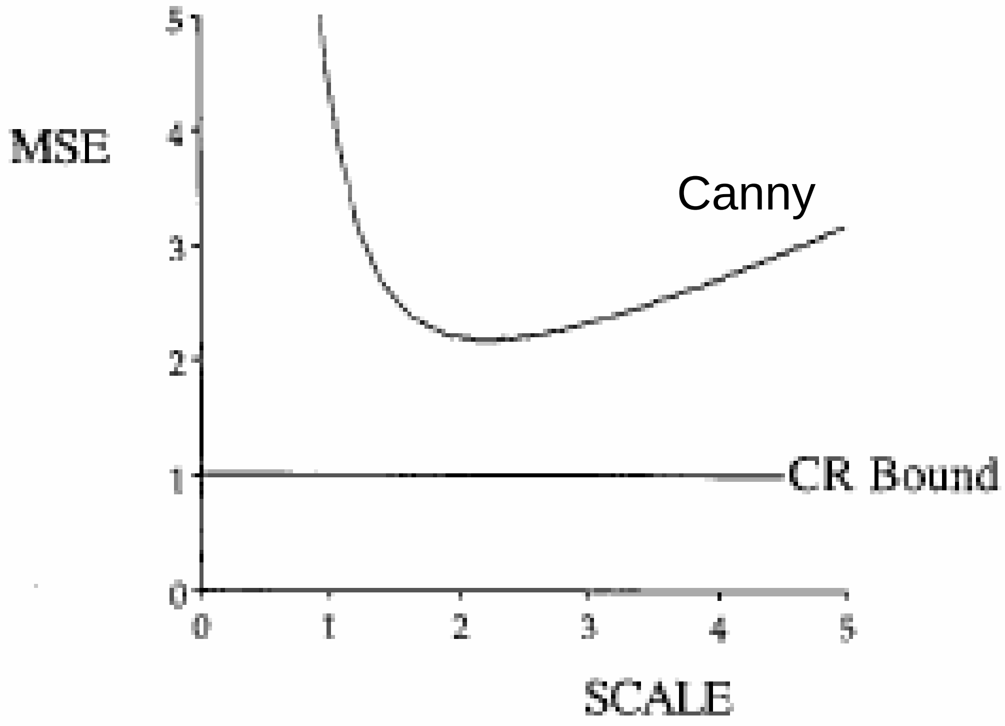


. 1. **Typical edge profile:** A plot of $S(x;\Theta) = I\Phi(\frac{x-\ell}{\sigma_s})$ along the x axis with $I = 2$, $\ell = 0$, and $\sigma_s = 0.6$.

"On achievable accuracy in edge localization" Kakarala and Hero, T. PAMI 14:7, 1992

Achievable accuracy in edge localization

$$E[(\hat{l} - l)^2] \geq \frac{\sigma^2 \sqrt{\pi} \sigma_s}{I^2 T}$$



Milanfar et al. EE Dept, UCSC

Overall Conclusions

- Fundamental performance limits in imaging are important to our understanding, and can help stop bickering.
- By understanding these limits, we optimize our algorithms accordingly.
- The propagation of these uncertainties from low to high level tasks is a challenge.
- Need to view image processing with increased awareness of notion of *information*.